

Otimização Inteira: Formulações e Relaxações

Prof. Alexandre Salles da Cunha

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
Belo Horizonte, Brasil

`acunha@dcc.ufmg.br`

Agosto 2020



UNIVERSIDADE FEDERAL
DE MINAS GERAIS



Relativas à esta apresentação

- 1 L. Wolsey. Integer Programming, Wiley, 1998. [Capítulos 1,2]

Considere o **Problema de Programação Linear (PPL)**

$$\max\{c^T x : Ax \leq b, x \geq 0\},$$

onde $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^n$.

- ❶ Se **algumas (não todas) variáveis de decisão** devem assumir **valores inteiros**, temos o *Problema de Programação Inteira Mista*:

$$(\text{MIP}) \max\{c^T x + h^T y : Ax + Gy \leq b, x \geq 0, y \text{ inteiro}\},$$

- ❷ Se **todas as variáveis de decisão** devem assumir **valores inteiros**, temos o *Programa Inteiro*:

$$(\text{IP}) \max\{c^T x : Ax \leq b, x \geq 0, x \text{ inteiro}\},$$

- ❸ Se **todas as variáveis de decisão** devem assumir **valores 0 ou 1**, temos o *Programa Binário*:

$$(\text{BIP}) \max\{c^T x : Ax \leq b, x \in \{0, 1\}^n\}.$$

- Dados:
 - **Conjunto base** $N = \{1, 2, \dots, n\}$.
 - **Pesos** $\{c_j : j \in N\}$ associados aos elementos do conjunto base.
 - Uma família $\mathcal{F}$ de subconjuntos de N (uma regra construtiva) que **define quais subconjuntos de N são válidos**.
- O **Problema de Otimização Combinatória** definido sobre $\mathcal{F}$ é:

$$(\text{COP}) \quad \min \left\{ \sum_{j \in S} c_j : S \in \mathcal{F} \right\}.$$

- 1 Muitas vezes, o mesmo problema de otimização pode ser formulado como um COP, MIP ou como um IP.
- 2 A forma como o problema é formulado (IP, MIP, ou como um COP) sugere o tipo de algoritmo que será empregado para resolvê-lo.

Qualidade de uma formulação

A *qualidade* da formulação empregada para se representar um problema de otimização em variáveis inteiras é de extrema importância para possam ser resolvidos. A definição de qualidade de uma formulação será apresentada em breve. *Essencialmente, depende dos limites duais que é capaz de fornecer...*

O processo de **formular ou modelar envolve as etapas:**

- Escolher variáveis de decisão, indentificando quais são contínuas, discretas, binárias.
- Representar o conjunto de restrições e função objetivo por meio de funções lineares, no caso de IPs ou MIPs.

Observações:

- Eventualmente, o modelo pode não ser linear. Neste caso temos um NLIP (Problema de Programação Inteira Não Linear).
Alguma técnica de linearização pode ser empregada ou, parte-se para resolver o NLIP (tema não coberto neste curso).
- O processo deve ser conduzido tendo-se em mente qual método de resolução é mais adequado para o tipo de formulação obtido.
- Possivelmente, haverá uma etapa posterior à modelagem inicial, dedicada a fortalecer a formulação (melhorando a sua qualidade, seus limites...).

- ① Problema do Caminho Mais Curto.
- ② Problema da Mochila $\{0 - 1\}$
- ③ Problema de Localização de Facilidades.
- ④ Problema de Cobertura de Conjuntos.
- ⑤ Problema do Caixeiro Viajante.
- ⑥ Problema de Sequenciamento de Tarefas.

Uma empresa dispõe de um orçamento de b unidades monetárias que serão empregadas para selecionar alguns de n projetos de investimento disponíveis. Para que o projeto j seja selecionado, é necessário realizar um investimento de a_j unidades monetárias. Neste caso, o projeto oferece um retorno de c_j unidades monetárias.

Objetivo: encontrar uma carteira de projetos que maximize o retorno e que não exceda o orçamento.

- **Variáveis de decisão binárias:** $x_j : j = 1, \dots, n$.

A variável x_j assume valor 1 caso o projeto j seja escolhido, 0 caso contrário.

- **Restrições:**

- O orçamento não deve ser excedido: $\sum_{j=1}^n a_j x_j \leq b$,

- Natureza bivalente das variáveis: $x_j \in \{0, 1\}$, $j = 1, \dots, n$

- **Função Objetivo: máximo retorno.** $\max \sum_{j=1}^n c_j x_j$.

$$\max \left\{ \sum_{j=1}^n c_j x_j : x \in P \cap \mathbb{B}^n \right\},$$

onde a região poliédrica P é definida por:

$$\sum_{j=1}^n a_j x_j \leq b.$$

$$0 \leq x_j \leq 1 \quad j = 1, \dots, n$$

- **Conjunto base:** $N = \{1, \dots, n\}$
- **Pesos** $\{c_j : j \in N\}$ associados aos elementos do conjunto base.
- **Família de subconjuntos** $\mathcal{F}$ (definem as mochilas, que são sistemas independentes)

$$\mathcal{F} = \{S \subseteq N : \sum_{j \in S} a_j \leq b\}$$

$$(\text{COP}) \quad \max \left\{ \sum_{j \in S} c_j : S \in \mathcal{F} \right\}$$

Um vendedor deve visitar um conjunto $N = \{1, \dots, n\}$ de clientes. O custo de deslocamento incorrido ao sair do cliente $i \in N$ diretamente para o cliente $j \in N \setminus \{i\}$ é c_{ij} . Estes custos podem ser assimétricos, isto é, $c_{ij} \neq c_{ji}$.

Objetivo: Definir uma rota ou uma sequência de clientes a serem visitados, de mínimo custo de deslocamento. Cada cliente deve ser visitado exatamente uma vez na rota.

Grafo direcionado

- $D = (N, A)$: O conjunto de vértices N representa os clientes. A , o conjunto de arcos, representa as conexões orientadas diretas entre eles.
- D pode ser completo ou esparso.
- $\delta^+(S) = \{(i, j) \in A : i \in S, j \notin S\}$
- $\delta^-(S) = \{(i, j) \in A : i \notin S, j \in S\}$
- $\delta^+(i), \delta^-(i)$: conjunto de arcos que partem e chegam de/em $i \in N$, respectivamente.

- **Variáveis de decisão binárias:** $x_{ij} : (i, j) \in A$, assumindo valor 1 caso o cliente j seja visitado imediatamente após o cliente i (o arco (i, j) é percorrido), 0, caso contrário.

- **Restrições:**

- Restrições de grau (de/para):

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1, i \in N$$

- Conexidade da solução (evitar sub-ciclos):

$$\sum_{(i,j) \in \delta^+(S)} x_{ij} \geq 1, \quad S \subset N, S \neq \emptyset,$$

- Natureza binária das variáveis: $x_{ij} \in \{0, 1\} : (i, j) \in A$.

- **Função Objetivo:** $\min \sum_{(i,j) \in A} c_{ij} x_{ij}$

$$\min \left\{ \sum_{(i,j) \in A} x_{ij} c_{ij} : x \in P_x \cap \mathbb{B}^{|A|} \right\}$$

onde P_x é definido por

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(S)} x_{ij} \geq 1, \forall S \subset N, S \neq \emptyset$$

$$x_{ij} \geq 0, (i,j) \in A$$

compacta = requer número polinomial (em $|A|, |N|$) de restrições e variáveis.

- ① Emprega outras restrições que proíbem sub-rotas (rotas envolvendo menos que de $|N|$ vértices).
- ② A formulação emprega as variáveis $\{x_{ij} : (i, j) \in A\}$ definidas anteriormente e um novo conjunto de variáveis inteiras,

$$\{t_i \in \{0, 1, \dots, |N| - 1\} : i \in N\}.$$

- ③ A variável t_i indica a ordem de visita do cliente $i \in N$, a partir de um ponto de partida (um vértice de N) que pode ser arbitrariamente escolhido. Vamos designar o ponto de partida como $r \in N$, de forma que $t_r = 0$.

- 1 Devemos garantir que se o vértice $j \neq r$ é visitado imediatamente após o vértice i (e neste caso, $x_{ij} = 1$), então vale:

$$t_j \geq t_i + 1.$$

Caso contrário, isto é, se $x_{ij} = 0$ e portanto j não é visitado imediatamente após i , esta restrição não deve ser imposta.

- 2 Para escrever as restrições que modelam este par de situações, vamos lembrar que $t_i \in \{0, 1, \dots, |N| - 1\}$.
- 3 Vamos empregar uma **disjunção para representar este fato**:

$$t_j \geq t_i + 1 - (1 - x_{ij})|N|, \text{ para todo } (i, j) \in A, j \neq r$$

- 4 Observe que quando $x_{ij} = 0$, a restrição acima é trivialmente satisfeita. Por outro lado, quando $x_{ij} = 1$, obtemos exatamente o que desejávamos.

$$\min \left\{ \sum_{(i,j) \in A} x_{ij} c_{ij} : (x, t) \in P_{xt} \cap (\mathbb{B}^{|A|} \times \mathbb{Z}^{|N|}) \right\}$$

onde P_{xt} é definido por

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1, i \in N$$

$$t_j \geq t_i + 1 - (1 - x_{ij})|N|, \text{ para todo } (i,j) \in A, j \neq r$$

$$x_{ij} \geq 0, (i,j) \in A$$

$$t_r = 0, t_i \in [1, |N| - 1], i \in N \setminus \{r\}$$

- ❶ Qualidade dos *limites duais* associados à formulação. Devemos verificar qual formulação fornece limites duais mais fortes, isto é, que para qualquer vetor de custos sempre fornece o valor ótimo da relaxação contínua pelo menos tão elevado quanto as demais.
- ❷ O esforço computacional necessário para avaliar estes limites também é um aspecto importante.
- ❸ No caso das formulações discutidas para o TSP:
 - Formulações P_x, P_{xt} são definidas em espaços diferentes de variáveis, com um conjunto comum, x .
 - Para mostrar que P_x é melhor que P_{xt} devemos mostrar que $P_x \subseteq \text{Proj}_x(P_{xt})$, onde $\text{Proj}_x(P_{xt})$ denota a projeção de P_{xt} no espaço x e que existe $(\hat{x}, \hat{t}) \in P_{xt} : \hat{x} \notin P_x$.

O problema consiste em definir um plano de produção de mínimo custo, para um único produto, em um período de tempo de n períodos. Além do estoque inicial do produto, s_0 , os seguintes dados são conhecidos:

- $f_t \geq 0$: custo fixo incorrido caso haja produção no período t ,
- $p_t \geq 0$: custo unitário de produção em t ,
- $h_t \geq 0$: custo unitário de estoque em t ,
- d_t : demanda a ser satisfeita em t .

- **Variáveis de decisão:**

- $x_t \geq 0$: quantidade produzida em t ,
- $s_t \geq 0$: quantidade de estoque ao final de t ,
- $y_t \in \{0, 1\}$, assumindo valor 1 se a produção em t for permitida, 0, c.c.

- **Dois aspectos são importantes na modelagem:**

- Garantir o balanço de fluxo em cada instante de tempo t .
- Garantir que o custo fixo seja incorrido em t se houver produção $x_t > 0$. Isto é, $x_t > 0 \rightarrow y_t = 1$.

- Vamos assumir que um **limite superior** M (por exemplo $M = \sum_{t=1}^n d_t$, se $s_n = 0$) na quantidade a ser produzida em cada período de tempo t seja conhecido.

$$\min \left\{ \sum_{t=1}^n p_t x_t + \sum_{t=1}^n h_t s_t + \sum_{t=1}^n f_t y_t : (x, s, y) \in P_1 \cap (\mathbb{R}^n \times \mathbb{R}^{n+1} \times \mathbb{B}^n) \right\}$$

onde P_1 é a região poliédrica:

$$s_{t-1} + x_t = d_t + s_t, \quad t = 1, \dots, n,$$

$$x_t \leq M y_t, \quad t = 1, \dots, n,$$

$$s_0 = 0,$$

$$s_t, x_t \geq 0,$$

$$0 \leq y_t \leq 1, \quad t = 1, \dots, n.$$

- Observe que $s_t = s_0 + \sum_{i=1}^t x_i - \sum_{i=1}^t d_i, t \geq 1$
- Vamos assumir $s_0 = 0$ e re-escrever a função objetivo:

$$\begin{aligned}
 & \sum_{t=1}^n p_t x_t + \sum_{t=1}^n h_t s_t + \sum_{t=1}^n f_t y_t = \\
 & \sum_{t=1}^n p_t x_t + \sum_{t=1}^n h_t \left(\sum_{i=1}^t x_i - \sum_{i=1}^t d_i \right) + \sum_{t=1}^n f_t y_t = \\
 & \sum_{t=1}^n \left(p_t + \sum_{i=t}^n h_i \right) x_t - \sum_{t=1}^n h_t \sum_{i=1}^t d_i + \sum_{t=1}^n f_t y_t = \\
 & \sum_{t=1}^n c_t x_t + \sum_{t=1}^n f_t y_t - K,
 \end{aligned}$$

onde $K = \sum_{t=1}^n h_t \sum_{i=1}^t d_i$ e $c_t = p_t + \sum_{i=t}^n h_i$.

- **Variáveis de decisão empregadas:**

- $\{w_{it} \geq 0, i = 1, \dots, n, t = i, \dots, n\}$. w_{it} é a quantidade produzida no período i para atender a demanda do período t .
- $\{y_t \in \{0, 1\}, t = 1, \dots, n\}$: definida anteriormente.

- **Novas restrições:**

$$\sum_{i=1}^t w_{it} = d_t, \quad t = 1, \dots, n,$$

$$w_{it} \leq d_t y_i, \quad i = 1, \dots, n, t = i, \dots, n$$

$$w_{it} \geq 0, \quad i = 1, \dots, n, t = i, \dots, n$$

$$y_t \in \{0, 1\} \quad t = 1, \dots, n$$

- **Função Objetivo:**

$$\min \sum_{i=1}^n \sum_{t=i}^n c_i w_{it} + \sum_{t=1}^n f_t y_t$$

- ❶ Considere a seguinte solução fracionária:

$$x_t = d_t, y_t = \frac{d_t}{M}, \quad t = 1, \dots, n.$$

- ❷ Observe que ela satisfaz as restrições que definem P_1 :

$$s_{t-1} + x_t = d_t + s_t \Rightarrow s_t = 0, \quad t = 1, \dots, n$$

$$x_t \leq M y_t \Rightarrow d_t \leq M \frac{d_t}{M}, \quad t = 1, \dots, n$$

$$x_t \geq 0, 0 \leq y_t \leq 1, \quad t = 1, \dots, n$$

- ❸ Porém, **não satisfaz as restrições que definem a formulação P_2 :**

$$w_{ii} = d_i, w_{it} = 0, \quad i = 1, \dots, n, t > i$$

$$w_{it} \leq d_t y_i, \quad i = 1, \dots, n, i \leq t \Rightarrow$$

quando $i = t$ temos: $d_i = w_{ii} \leq d_i y_i = d_i \left(\frac{d_i}{M} \right) < d_i$

Como provamos que o problema

$$z = \{\max c^T x : x \in X \subseteq \mathbb{Z}^n\}$$

foi resolvido ? Resolver o problema consiste em:

- 1 Obter uma solução viável $x^* \in X$ e
- 2 Garantir que $c^T x \leq c^T x^*$ para qualquer $x \in X$.

Precisamos encontrar:

- Uma solução viável x^* que (no caso de max) fornece um limite inferior válido para o valor ótimo: $z \geq \underline{z} = c^T x^*$;
- Um limite superior válido $\bar{z}$, denominado limite dual: $z \leq \bar{z}$.

Se estes limites coincidirem, $\bar{z} = \underline{z}$, provamos que x^* resolve o problema.

- Gerar uma sequência não crescente (idealmente decrescente) de limites superiores $\{\bar{z}^k\}$ válidos
 - Gerar uma sequência não decrescente (idealmente crescente) de limites inferiores $\{\underline{z}^k\}$
 - Quando $\bar{z}^k = \underline{z}^k$ para algum k , o algoritmo pára, com um certificado de otimalidade para a melhor solução viável obtida.
-
- Se o critério de parada é: $\bar{z}^k = \underline{z}^k$, para algum k , a solução encontrada é ótima, e o algoritmo é exato.
 - Se o critério de parada é: $\frac{\bar{z}^k - \underline{z}^k}{\underline{z}^k} \leq \varepsilon$, a solução é dita ε -ótima. Trata-se de um algoritmo com garantia de distância.

- Algum procedimento para obter soluções viáveis, por exemplo, Heurísticas.
- Algum procedimento para obter limites duais. Um dos principais mecanismos é via *Relaxações*.

A idéia central consiste em substituir o problema original, geralmente *difícil* de ser resolvido, por um problema *mais fácil*, cuja função objetivo ótima é não menor que z (caso maximização).

Como definir este problema relaxado:

- 1 Ampliando o conjunto de soluções viáveis, de forma que se otimize sobre um conjunto maior.
- 2 Substituindo a função objetivo por outra que assuma um valor não menor que a original em qualquer solução no domínio do problema original.

Definição

Dizemos que um problema de otimização (RP)

$$z^R = \max \{f(x) : x \in T \subseteq \mathbb{R}^n\}$$

é uma relaxação para o (IP)

$$z = \max \{c^T x : x \in X \subseteq \mathbb{R}^n\}$$

se as duas condições seguintes são satisfeitas:

- ① $X \subseteq T$ (o conjunto de soluções viáveis do problema relaxado contém o conjunto de soluções original).
- ② $f(x) \geq c^T x$ para qualquer $x \in X$.

Proposição

Se RP é uma relaxação para IP, então $z^R \geq z$

Demonstração

- Seja x^* a solução de IP.
- Por um lado, como $x^* \in X$ e $X \subseteq T$ temos que $x^* \in T$.
- Por outro, temos que $c^T x^* \leq f(x^*)$.
- Pela otimalidade de z^R , temos então que $z^R \geq f(x^*) \geq c^T x^* = z$.



- 1 Relaxação de Programação Linear.
- 2 Relaxações Combinatórias.
- 3 Relaxação Lagrangeana.

Definição - Relaxação de Programação Linear

Dado o IP

$$z = \{\max c^T x : x \in P \cap \mathbb{Z}^n\}$$

e a formulação

$$P = \{x \in \mathbb{R}^n : Ax \leq b\},$$

a **Relaxação de Programação Linear** associada ao IP corresponde ao Programa Linear

$$z^{LP} = \{\max c^T x : x \in P\}.$$

- A função objetivo de RP é não menor que a de IP, para qualquer solução viável do IP.
- $P \cap \mathbb{Z}^n \subseteq P$: o domínio de RP contém o domínio de IP.
- Logo $z \leq z^{LP}$.

Exemplo

Dado o IP:

$$\begin{aligned}
 z &= \max & 4x_1 & -x_2 \\
 & & 7x_1 & -2x_2 \leq 14 \\
 & & & x_2 \leq 3 \\
 & & 2x_1 & -2x_2 \leq 3 \\
 & & x & \in \mathbb{Z}_+^2
 \end{aligned}$$

- O vetor $(2, 1)^T$ é uma solução viável, logo $z \geq 7$.
- Para a relaxação linear, temos $z^{LP} = \frac{59}{7}$, obtida para $\bar{x} = (\frac{20}{7}, 3)^T$.
- Como os coeficientes da função objetivo do IP são inteiros temos:

$$z \leq \lfloor \frac{59}{7} \rfloor = 8.$$

A qualidade dos limites duais determina a qualidade da formulação

Dadas duas formulações P_1, P_2 para o

$$(IP) \quad z = \{\max c^T x : x \in X \subseteq \mathbb{Z}^n\},$$

tais que P_1 é uma formulação melhor que P_2 , isto é: $P_1 \subset P_2$.

Se $z_i^{LP} = \{\max c^T x : x \in P_i\}$ para $i = 1, 2$ denota o valor das respectivas relaxações lineares, então $z_1^{LP} \leq z_2^{LP}$ para todo $c \in \mathbb{R}^n$.

Proposicao:

Se RP é inviável, então o IP é inviável.

Demonstração:

Seja $X = \emptyset$ o conjunto de soluções do problema relaxado (RP) e T o conjunto de soluções do problema original (IP). Como RP é uma relaxação temos que $T \subseteq X = \emptyset$, logo $T = \emptyset$. □

Proposicao:

Se a solução $\bar{x}$ para o problema relaxado RP é viável para (IP) e se $f(\bar{x}) = c^T \bar{x}$, então $\bar{x}$ resolve IP.

Demonstração:

Como $\bar{x}$ é uma solução viável, $z \geq c^T \bar{x}$. Por outro lado, $f(\bar{x})$ fornece um limite superior válido para z . Como por hipótese $f(\bar{x}) = c^T \bar{x}$, temos $z \leq f(\bar{x}) = c^T \bar{x} \geq z$ e então $z = f(\bar{x})$. □

Exemplo

Uma solução ótima da Relaxação Linear do IP

$$\begin{aligned} z = \max \quad & 7x_1 + 4x_2 + 5x_3 + 2x_4 \\ & 3x_1 + 3x_2 + 4x_3 + 2x_4 \leq 6 \\ & x \in \mathbb{B}^4 \end{aligned}$$

é dada por $\bar{x} = (1, 1, 0, 0)^t \in \mathbb{B}^4$.

Como $\bar{x}$ satisfaz o requisito de integralidade, $\bar{x}$ resolve o IP.

Definição

Sempre que o problema relaxado for um problema de otimização combinatória, a relaxação associada é dita *combinatória*.

Exemplos:

- Caixeiro Viajante Assimétrico: Problema de Atribuição.
- Problema da Mochila: Problema da Mochila em Coeficientes Inteiros.

Entrada: Digrafo $D = (N, A)$, custos associados aos arcos $\{c_{ij} : (i, j) \in A\}$.

$$\min \left\{ \sum_{(i,j) \in A} x_{ij} c_{ij} : x \in P_x \cap \mathbb{B}^{|A|} \right\}$$

onde P_x é definido por

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(S)} x_{ij} \geq 1, \forall S \subset N, S \neq \emptyset$$

$$x_{ij} \geq 0, (i, j) \in A$$

$$z^{atrib} = \min \left\{ \sum_{(i,j) \in A} x_{ij} c_{ij} : x \in P_{atrib} \cap \mathbb{B}^{|A|} \right\}$$

onde P_{atrib} é dada por:

$$\sum_{(j,i) \in \delta^-(i)} x_{ji} = 1, i \in N$$

$$\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1, i \in N$$

$$x_{ij} \geq 0, (i,j) \in A$$

Observe que:

- $P_x \subset P_{atrib}$ (isto vale para qualquer formulação P para o TSP).
- $z \geq z^{atrib}$.

Proposição

Uma relaxação do conjunto $X = \{x \in \mathbb{Z}_+^n : \sum_{j=1}^n a_j x_j \leq b\}$ é dada pelo conjunto $X' = \{x \in \mathbb{Z}_+^n : \sum_{j=1}^n \lfloor a_j \rfloor x_j \leq \lfloor b \rfloor\}$.

Prova

$$\sum_{j=1}^n \lfloor a_j \rfloor x_j \leq \sum_{j=1}^n a_j x_j \leq b$$

Como os vetores de incidência $x \in X$ são inteiros, temos que $\sum_{j=1}^n \lfloor a_j \rfloor x_j$ deve ser necessariamente uma quantidade inteira. Logo $\sum_{j=1}^n \lfloor a_j \rfloor x_j \leq \lfloor b \rfloor$. Ou seja: $x \in X \rightarrow x \in X'$. □

Vamos considerar o IP

$$z = \{\max c^T x : Ax \leq b, x \in S \subseteq \mathbb{Z}^n\}.$$

- Podemos *relaxar* um conjunto de restrições, por exemplo $\{Ax \leq b\}$ e resolver a relaxação

$$z^R = \{\max c^T x : x \in S \subseteq \mathbb{Z}^n\}.$$

- Claramente $z \leq z^R$.
- Na Relaxação Lagrangeana, não apenas relaxamos um conjunto de restrições que dificultam a solução de um IP. Também associamos a cada restrição i relaxada um multiplicador de Lagrange u_i e somamos o produto da folga $(b_i - a_i^T x)$ da restrição pelo multiplicador na função objetivo da relaxação.

Proposição

Seja (PLAG)

$$z(u) = \max \{c^T x + u^T (b - Ax) : x \in S\}.$$

Então $z(u) \geq z, \forall u \geq 0$.

Demonstração

Seja $\bar{x}$ a solução ótima do IP.

- Como $\bar{x}$ resolve o IP, $\bar{x} \in S$ é uma solução viável para (PLAG).
- Como $u \geq 0$ e $A\bar{x} \leq b$ temos que $z = c^T \bar{x} \leq c^T \bar{x} + u(b - A\bar{x})$.
- Por outro lado, pela otimalidade de $z(u)$ temos que $z(u) \geq c^T \bar{x} + u(b - A\bar{x})$.
- Logo $z = c^T \bar{x} \leq c^T \bar{x} + u(b - A\bar{x}) \leq z(u)$. □

Considere o par de problemas

$$(P) \quad \{ \max c^T x : Ax \leq b, x \in \mathbb{R}_+^n \}$$

$$(D) \quad \{ \min u^T b : u^T A \geq c^T, u \in \mathbb{R}_+^m \}$$

- Sabemos que $c^T x \leq u^T b$ para quaisquer vetores $\{x \in \mathbb{R}_+^n : Ax \leq b\}$ e $\{u \in \mathbb{R}_+^m : u^T A \geq c\}$.
- Em particular, se $\bar{x}$ e $\bar{u}$ resolvem P e D, respectivamente, temos que $c^T \bar{x} = \bar{u}^T b$ (Dualidade forte).

Definição

Os problemas

$$(P) \quad z = \{ \max c(x) : x \in X \}$$

$$(D) \quad w = \{ \min w(u) : u \in U \}$$

formam um par primal-dual fraco se $c(x) \leq w(u)$ para quaisquer $x \in X, u \in U$. Se, além disto $z = w$ então (P) e (D) formam um par primal-dual forte.

Vantagem da dualidade sobre a relaxação

A relaxação só fornece um limite superior válido se resolvida na otimalidade. Por outro lado, qualquer solução viável de D fornece um limite superior válido para P.

Proposição

Os problemas (P) $z = \{\max c^T x : Ax \leq b, x \in \mathbb{Z}_+^n\}$ e (D) $w^{LP} = \{\min u^T b : u^T A \geq c^T, u \in \mathbb{R}_+^m\}$ formam um par primal-dual fraco.

Demonstração

- Considere (RLP) $z^{LP} = \{\max c^T x : Ax \leq b, x \in \mathbb{R}_+^n\}$.
- Claramente $z \leq z^{LP}$.
- Como (RLP) e (D) formam um par primal-dual, temos que $z^{LP} \leq w^{LP}$.
- Logo $z \leq z^{LP} \leq w^{LP}$. □

Emparelhamento ou *matching*

Dado um grafo $G = (V, E)$ um emparelhamento (*matching*) consiste em um conjunto de arestas $M \subseteq E$ tal que não existem duas arestas em M que possuem a mesma extremidade.

Cobertura de arestas por vértices

Dado um grafo $G = (V, E)$ uma cobertura das arestas de E por vértices é um conjunto $R \subset V$ tal que toda aresta de E possui uma extremidade em R .

Proposição

Os problemas (P) $\{\max_{M \subseteq E} |M| : M \text{ é um emparelhamento de } E\}$ e (D) $\{\min_{C \subseteq V} |C| : C \text{ é uma cobertura de arestas por vértices}\}$ formam um par primal-dual fraco.

Prova

- Definimos $A \in \mathbb{B}^{n \times m}$ ($n = |V|$, $m = |E|$) tal que $a_{j,e} = 1$ se $j \in V$ é extremidade de $e \in E$ (0, caso contrário) e escrevemos (P) e (D) como $z_{\text{match}} = \left\{ \max \sum_{j=1}^m x_j : Ax \leq 1, x \in \mathbb{Z}_+^m \right\}$ e $z_{\text{cov}} = \left\{ \min \sum_{i=1}^n y_i : y^T A \geq 1, y \in \mathbb{Z}_+^n \right\}$.
- Sejam $z_{\text{match}}^{\text{PL}}$ e $z_{\text{cov}}^{\text{PL}}$ os valores ótimos das respectivas relaxações lineares.
- Então: $z_{\text{cov}}^{\text{PL}} \leq z_{\text{cov}}$ e $z_{\text{match}} \leq z_{\text{match}}^{\text{PL}}$. Como $z_{\text{match}}^{\text{PL}} \leq z_{\text{cov}}^{\text{PL}}$, temos $z_{\text{match}} \leq z_{\text{match}}^{\text{PL}} \leq z_{\text{cov}}^{\text{PL}} \leq z_{\text{cov}}$. □